SCF Tabling Utility Instructions

Introduction

The SCF tabling utility contains of a set of macros, written in Visual Basic code, that allow users to create a table of weighted descriptive statistics applying Excel to an Extract file of SCF data, which is documented separately. This utility allows users to output one of three descriptive statistics: weighted median, weighted mean, or weighted percent. These statistics are weighted by an analysis weight. Excel maintains built-in worksheet functions that calculate medians, means, and frequencies (percents). However, because the SCF sample is not an equal-probability design, weights play a critical role in interpreting the survey data and calculating descriptive statistics. These weights are intended to compensate for unequal probabilities of selection in the original design and for unit nonresponse (failure to obtain an interview). For a detailed discussion of these weights, see "Consistent Weight Design for the 1989, 1992, and 1995 SCFs and the Distribution of Wealth," by Arthur B. Kennickell and R. Louise Woodburn. The purpose of this utility is to allow users to perform weighted mean and weighted percent calculations efficiently and to provide users the capacity to output weighted medians in Excel. This utility is motivated by a desire to increase the accessibility of SCF data to users. This software was used to produce the tables included in the January 2003 *Federal Reserve Bulletin* article, "Recent Changes in U.S. Family Finances: Evidence from the 1998 and 2001 Survey of Consumer Finances". Although this software has been tested extensively errors still remain; the Federal Reserve does not take responsibility for the use analysts make of this software.

Opening the SCF extract that includes the SCF tabling utility

The SCF website includes two data extracts. One extract includes only the aggregate data (e.g. scfp2001.xls); this version is made available to users who do not wish to use the tabling utility. The other extract includes both the aggregate data and the SCF tabling utility (e.g. SCFTabling2001.xls). To begin using the utility download the second file and open it in Excel. This file contains macros. When the file opens, a dialog box warns users that macros may potentially corrupt the workbook and provides users the option to either enable or disable macros. In order to have access to the SCF tabling utility users must click "Enable Macros".

How to make the SCF tabling utility form appear

Opening the file that includes the tabling macro will add three menu options to Excel's toolbars, which provide users access to the utility. First, a button with the appearance of a table appears at the end of the command bar. When the cursor is placed over this button a message appears that says, "SCF Tabling". Second, a menu item named "SCF Tabling" is appended to the bottom of the tools menu. Finally, an "SCF Tabling" menu item is added to the bottom of the right-click shortcut menu. To find this item simply right-click anywhere on a worksheet. SCF Tabling is located at the bottom of the shortcut menu that appears. When users click on any one of these menu items the SCF tabling utility appears. As noted in more detail later, when Excel closes normally all of these menu items disappear until the next time this file is opened.

Using the SCF tabling utility

The SCF tabling utility has many options to choose from. Users can select one of three descriptive statistics: weighted median, weighted mean, or weighted percent, all of which may be classified by as many as three conditions at a time. Users can write the requested table to an existing worksheet or to entirely new worksheet. The SCF Tabling form is structured with seven frames of options. Once users make all of their selections from these frames and click "OK" the macros calculate all of the requested statistics and place them in a table in the specified worksheet. Please note that each calculation takes a few seconds and computing speed varies by machine. Large tables may take several minutes to produce. The following sections describe each frame's function.

Frame 1: Output Worksheet

This frame allows users to indicate where to store the requested table. Users can either select to output their table to an existing worksheet or to an entirely new worksheet. To output a table to a new worksheet click the button next to "New worksheet" and enter a valid Excel worksheet name in the box with the header "Type or select worksheet name". Valid Excel worksheet names may not exceed 31 characters or include any of the following symbols: \ / ? : *. Users may also output their table to an existing worksheet. To output a table to an existing worksheet simply select the button next to "Existing worksheet". All valid worksheet names (note: tables can not be output to the worksheet that holds the data) are included in the drop down box with the heading "Type or select worksheet name". Click the arrow on the right side of this box and select the desired output worksheet from the list. The requested table appears at the bottom of the selected worksheet.

Frame 2: Weighted Statistics

The SCF tabling utility allows users to select one of three statistics: weighted median, weighted mean, or weighted percent. It should be noted that the public version of the SCF data are released without any of the data or weight adjustments that are applied in creating the tables in the *Bulletin* article. Consequently, the requested output – particularly weighted means– in some instances may be strongly affected by outliers. Further, many of the aggregate variables included in this data set have nonnormal distributions. For variables with nonnormal distributions the weighted median is a better measure of central tendency than the weighted mean, which is affected by outliers.

Weighted median:

The weighted median corresponds to the data point that lies at the 50th percentile of the ordered weighted distribution.

Weighted mean:

The weighted mean is the summation of the product of the variable and the analysis weight divided by the summation of the analysis weight,

$$\sum_{i=1}^{n} X_i * WGT_i \,/\, \sum_{i=1}^{n} WGT_i$$

where X is the variable of interest, WGT is the analysis weight, and i denotes the observation number.

Weighted percent:

A weighted percent is the ratio of the summation of weights for those families with holdings of the variable divided by the summation of all weights,

$$\sum_{i=1}^{n} I(X_i \neq 0) * WGT_i \,/\, \sum_{i=1}^{n} WGT_i$$

where I is an indicator function that is set to 1 if X is not equal to zero and is set to zero otherwise.

Users select one of these statistics. The default statistic is weighted median.

Frame 3: All families

This frame allows users to include "all families" in their calculations or condition their output to consider only those families with holdings of the variable of interest. The statistics included in the *Bulletin* article are based on all families with holdings with the exception of the calculations where the aggregate variables INCOME, NETWORTH, or KGTOTAL are the variables of interest. For these variables it is assumed that all families have holdings. Some variables in this file are sparsely held by only a small number of families. It is worth noting that the weighted median is zero for any variable where less than fifty percent of families have holdings when "all families" are included in the calculation. To include "all families" in the requested calculations users should click the button next to "All families" in this frame. To include only families with holdings users should select the button next to "families with holdings". The default setting is "families with holdings". It should be noted that weighted percent calculations always include all families (by definition) even if "families with holdings" is selected.

Frame 4: Base-year dollars

The dollar values in each extract file are given in current dollars. Thus, all dollar-denominated variables in the 1998 extract file are given in 1998 dollars. The default settings if the SCF tabling utility produce dollar estimates in current dollars. The utility provides users the option to calculate their tables in real dollars of any one of the five survey years. The calculations included in the *Bulletin* article were computed in 2001 dollars. If users wish to make estimates comparable to those in the article for 1998, for example, they would click the button next to "2001" in frame 4; this would adjust the output from 1998 dollars to 2001 dollars.

Frame 5: Variables

Users must select at least one variable from the box in frame 5 to base their calculations on. This box lists all of the aggregate variables in the extract. User-defined variables also appear in this box and may be used in calculations. Users may select as many of these variables as they would like to include in the output table. Each selected variable is printed across the top of the output table in alphabetical order.

Frame 6: Conditioning variables and conditions

The SCF tabling utility allows users to condition their output by categories used in the *Bulletin* article or by the quantiles of any aggregate variable. Other conditioning may be made by specifying a particular population in frame 3 or providing a condition in frame 7.

There are eight *Bulletin* category groups to choose from: all families, percentile of income, age of head, education of head, race or ethnicity of respondent, current work status of head, housing status, and percentile of net worth.

The "All families" group (ALL) includes all families in the population specified by frame 3.

Percentile of income includes six categories:

INCCAT1: Less than 20
INCCAT2: 20-39.9
INCCAT3: 40-59.9
INCCAT4: 60-79.9
INCCAT5: 80-89.9
INCCAT6: 90-100

Age of head includes six categories:

AGECL1: Less than 35 years old
AGECL2: 35-44 years old
AGECL3: 45-54 years old
AGECL4: 55-64 years old
AGECL5: 65-74 years old
AGECL6: 75 or more years old

Education of head includes four categories:

EDCL1: No high school diploma
EDCL2: High school diploma
EDCL3: Some college
EDCL4: College degree

Race or ethnicity of respondent includes two categories:

RACECL1: White non-Hispanic
RACECL2: Nonwhite or Hispanic

Current work status of head includes four categories:

OCCAT11: Working for someone else
OCCAT12: Self-employed
OCCAT13: Retired
OCCAT14: Other not working

Housing status includes two categories:

HOUSECL1: Owner
HOUSECL2: Renter or other

Percentile of net worth includes five categories:

NWCAT1: Less than 25
NWCAT2: 25-49.9
NWCAT3: 50-74.9
NWCAT4: 75-89.9
NWCAT5: 90-100

These are the categories included in the *Bulletin* article. For a further description of these categories please refer to the article.

To condition output by one or more of these categories select "Categories" from the box with the heading "Select quantiles or categories" at the top of frame 5. Once "categories" is selected, the box on the left side of this frame is populated with each of these categories. Select the categories to condition by from this list. Users may select as many of these categories as they wish. These categories appear in the first column of the output table. The intersection of the variable at the top of the table with the category represents the conditional weighted statistic of the variable of interest conditioned by the specified category.

Users may condition their output by the deciles, quintiles, or quartiles of any of the aggregate variables. If user do so, quantiles must be selected from the box with the heading "Select quantiles or categories". Once quantiles is selected two things occur. First, the box on the left side of this frame is filled with a list of the aggregate variables. Users must select the variables to condition by from this list. Second, the box with the heading "Quantile type" lists quartiles, quintiles, and deciles. Users must select the quantile type from this list. Once one of these quantiles is selected, the box in the bottom right corner of this frame with the heading "Specify quantiles" lists the specific quantiles for the specified quantile type. Select specific quantiles from this box to include in the output table.

Frame7: User-specific condition

This frame allows users to condition their table by one user-specific condition. This frame includes a list of all of the aggregate variables, two text boxesm, and a check box. The user may select one aggregate variable from the list of variables to condition their output table. The first text box is for a lower bound specification. The second text box is for an upper bound specification. The default for the lower bound text box and the upper bound text box is the minimum value and maximum value of the selected conditioning variable, respectively. If the variable selected is a dollar variable then the check box must be selected. This allows the dollar value bounds set in the text boxes to be adjusted to the base-year selected in frame 4. This assumes that the limits set are given in base-year dollars. For example, to return descriptive statistics for families with income greater than $80,000, INCOME would be selected from the list of variables, 80000 would be entered into the lower bound text box, and the check box would be clicked. Please note: when a condition is set in this frame every output statistic in the corresponding table is subject to this condition.

Example 1:

Here are instructions to write a table with weighted medians in current dollars to a new worksheet "Table1" for "families with holdings" for the variables DEBT, CCBAL, and INSTALL, conditioned by *Bulletin* categories INCCAT1 through INCCAT6 and EDCL1 through EDCL4.

Frame 1: Type Table1 into the box with the heading "Type or select worksheet name"
Frame 2: Nothing. Weighted median is the default
Frame 3: Nothing. Families with holdings is the default.
Frame 4: Nothing in this example.
Frame 5: Select DEBT, CCBAL, and INSTALL from the list of variables. To select multiple variables hold down the ctrl key and click the variable names.
Frame 6: Select categories from the box with the heading "Select quantiles or categories". Then select INCCAT1 through INCCAT6 and EDCL1 through EDCL4 from the box on the left.
Frame 7: Nothing.
Click OK

Example 2:

Here are instructions to write a table with weighted means to an existing worksheet "deciles_of_asset" for all families for the variables INCOME and NETWORTH, conditioned by the deciles of ASSET.

Frame 1: Click the button for existing worksheet and then select "deciles_of_asset" from the list of existing worksheets in the box with the heading "Type or select worksheet name".
Frame 2: Click the button next to weighted mean
Frame 3: Click the button next to all households
Frame 4: Nothing in this example
Frame 5: Select INCOME and NETWORTH from the list of variables.
Frame 6: Select quantiles from the box with the heading "Select quantiles or categories". Select ASSET from the list of variables in the box on the left. Select deciles from the box with the heading "Quantile type". Finally, highlight all ten deciles in the box with the heading "Specify quantiles" to include in the output table.
Frame 7: Nothing
Click OK

Potential questions and a few warnings:

Is it ok to have other Excel workbooks open while working with the SCF tabling utility?

While it is possible to have other workbooks open it is best to close all other Excel workbooks before opening the extract and using the SCF tabling utility. There are several reasons to close all other workbooks. First, the extract is a very large data set for Excel to manage. The less you have open the faster the utility will run. Further, the SCF tabling utility has gone through extensive testing before its release. However, it is a new product that still contains bugs. The more workbooks users have open while running this utility the greater the chance these bugs will appear.

What should be done if multiple SCF Tabling menu items appear?

The SCF Tabling menu items that appear in the tools menu, on your command bar, and the shortcut menu are placed there when this workbook's "open event" runs. These menu items are removed when the workbook's "close event" runs. If the workbook's "close event" does not run properly, for any reason, these menu items are

not removed. The next time this workbook is opened two sets of menu items appear. When this happens users will want to remove the extra set of menu items. In order to do this go to the tools menu, select macro, and then macros. When the list of macros appears select the macro DelSCFTablingBtn from the list and click run. Repeat these steps for the macro DelSCFTablingFromShortCut and the macro DelSCFTablingMI. This process removes the extra set of menu items.

What should be done if the SCF Tabling menu items do not appear or seem to suddenly vanish?

The SCF Tabling menu items are removed when this workbook's "close event" runs. The close event runs once users click close from the file menu or click the "x" in the upper right corner of the workbook. The close event prompts users to save their work. If users select cancel from this prompt the SCF menu items will have been removed while the extract remains open. If this happens, close the utility and reopen it.

What if macros are disabled?

When macros are disabled the extract data will appear but the SCF tabling utility will not be available.

Why does some of the output appear in bold?

Some cells in a table may appear in bold. This happens when calculations are based on fifty or fewer implicates (ten or fewer actual interviews). The bold font warns users that these calculations may be influenced by outliers.

When an unwanted table is deleted from a worksheet and then a new table is added to the worksheet the new table does not appear right after any other remaining tables on the worksheet. Why does this happen?

This is one of the few known bugs in this utility. Ongoing work to correct this problem is underway. For now just be aware that tables may appear at odd places in the worksheet when tables are deleted.

What if user-defined variables do not appear in the variable lists in the SCF tabling utility?

This could happen for a couple of reasons. First, make sure that these variables have a name and it appears in the first cell of the column that contains the variable. Also, make sure that these variables have the same number of rows as the other aggregate variables. Second, make sure these variables are contained in columns directly next to the aggregate data. In other words, be sure the data are contiguous (no empty columns separate the aggregate data from user-defined data). If both of these conditions are satisfied user-defined variables should appear in the SCF tabling utility's lists.

What if a Visual Basic run-time error occurs?

We hope this never happens but there is a good chance it will. If this happens please write down the error number and what steps you took to produce the error then press end. After pressing end please send us an email at scf@frb.gov.

Frame 4 should not be used when statistics are conditioned by user-defined classification variables that are based on dollar variables. This will result in erroneous output.

Please do not change either the workbook name or the name of the worksheet that holds the data. The macros that run the utility are dependent on these names.

The SCF tabling utility does not work on Macintosh operating systems.